

The space of logically consistent classical processes without causal order

Ämin Baumeler^{1,2} and Stefan Wolf^{1,2}

¹*Faculty of Informatics, Università della Svizzera italiana, Via G. Buffi 13, 6900 Lugano, Switzerland*

²*Facoltà indipendente di Gandria, Lunga scala, 6978 Gandria, Switzerland*

Classical correlations without predefined causal order arise from processes where parties manipulate random variables, and where the order of these interactions is not predefined. No assumption on the causal order of the parties is made, but the processes are restricted to be logically consistent under any choice of the parties' operations. It is known that for three parties or more, this set of processes is larger than the set of processes achievable in a predefined ordering of the parties. Here, we model all classical processes without predefined causal order geometrically and find that the set of such processes forms a polytope. Additionally, we model a smaller polytope — the deterministic-extrema polytope — where all extremal points represent deterministic processes. This polytope excludes probabilistic processes that must be — quite unnaturally — *fine-tuned*, because any variation of the weights in a decomposition into deterministic processes leads to a logical inconsistency.

I. MOTIVATION AND MAIN RESULT

An assumption often made in physical theories, sometimes implicitly, is the existence of a *global time*. In particular, quantum theory is formulated with time as an intrinsic parameter. If one relaxes this assumption by requiring *local* validity of some theory and *logical* consistency only, then a larger set of correlations can be obtained, called *correlations without predefined causal order*. The processes that lead to such correlations are called processes without predefined causal order. Two motivations to study such correlations are quantum gravity and quantum non-locality. Quantum gravity motivates this research in the sense that on the one hand, relativity is a *deterministic* theory equipped with a *dynamic* spacetime; on the other hand, quantum theory is a *probabilistic* theory embedded in a *fixed* spacetime. This suggests that quantum gravity is relaxed in both aspects, *i.e.*, it is a *probabilistic* theory equipped with a *dynamic* spacetime [1]. Quantum non-local correlations [2–4] motivate this study since the possibility of a satisfactory causal explanation [5] for such correlations is questionable [3, 6–13]. Dropping the notion of a global time or of an *a priori* spacetime — as has been suggested from different fields of research [14–23] — dissolves this paradox. This can be achieved by defining causal relations based on free randomness (see Figure 1) as opposed to defining free randomness based on causal relations [24, 25]. Such an approach gives a dynamic character to causality; causal connections are not predefined but are derived from the observed correlations.

Relaxations of quantum theory where the assumption of a global time is dropped have recently been studied widely [1, 26–45] (see Ref. [46] for a review). Our work follows the spirit of an operational quantum framework for such correlations developed by Oreshkov, Costa, and Brukner [31]. Some correlations appearing in their quantum framework — for two parties or more — cannot be simulated by assuming a predefined causal order of the parties. Such correlations are termed *non-causal*. Analogously to non-locality, non-causal correla-

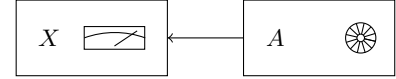


Figure 1. If the random variable A is an input (here, visualized by a knob), the random variable X is an output, and A is correlated to X , then A can signal to X which implies that X is in the *causal future* of A ($X \succeq A$).

tions could be witnessed by violating so-called causal inequalities [31, 32, 35, 43]. All causal inequalities in the two-party scenario and for binary inputs and outputs are presented in Ref. [43]. In a previous work [35], we showed that in the classical limit of the quantum framework, *i.e.*, if it is restricted to probability theory, *classical non-causal* correlations can arise as well. This result holds for three parties or more. In the present work we follow this path and give a *representation* of all classical — as opposed to quantum — processes without predefined causal order *as polytopes*. Such a representation helps in optimizing winning strategies for causal games [31, 43] — the optimization problem can be stated as a linear program —, and for finding new causal games.

First, we present the framework of classical correlations without predefined causal order. Then, we describe the polytope of processes that lead to such correlations implicitly and explicitly for scenarios with up to three parties and binary inputs and outputs. In the general case, we give an implicit description of the polytope. In addition, we construct the smaller polytope of classical processes without predefined causal order where all extremal points describe deterministic processes. We call this polytope the *deterministic-extrema polytope*. The processes from this polytope can be thought of as being “more physical” in the sense that its extremal points are not *proper mixtures* of logically inconsistent processes [47], *i.e.*, this set contains processes that can be written as a convex combination of deterministic ones from *within* the polytope only. Our motivation for this is that some proper mixtures need to be *fine-tuned* [13], *i.e.*, tiny variations of the mixtures renders the processes

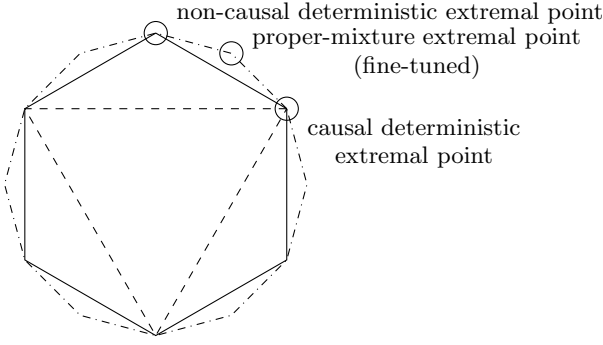


Figure 2. A qualitative representation of processes without predefined causal order studied in this work is given. The dashed region describes all processes that are achievable in a predefined causal order — it also forms a polytope [42, 43]. The polytope with the dashed-dotted lines is the polytope of processes without predefined causal order. The region in-between marked with the solid lines is the polytope of processes without predefined causal order restricted to deterministic extremal points.

logically inconsistent. The *fine-tuned* proper mixtures are the probabilistic extremal points of the larger polytope. A qualitative representation of these polytopes is given in Figure 2.

II. MODELLING CLASSICAL CORRELATIONS WITHOUT PREDEFINED CAUSAL ORDER

A. Causality, predefined causal order, and a framework of classical correlations without predefined causal order

We describe an operational framework without global assumptions (other than logical consistency). Causal relations are defined as in the interventionists' approach to causality [48, 49]: Outputs can be correlated to inputs and inputs are manipulated *freely* (see Figure 1). Defining causality based on *free randomness* is the converse approach to the one used in recent literature [24, 25]; there, free randomness is defined based on *causal relations*.

Definition 1 (Causality [35]). For two correlated random variables X and A , where X is an output and A is an input, *i.e.*, A is chosen *freely*, we say that X is in the *causal future* of A , or equivalently, that A is in the *causal past* of X , denoted by $X \succeq A$ or $A \preceq X$. The negations of these relations are denoted by $\not\succeq$ and $\not\preceq$.

Consider N parties $\{S_j\}_{0 \leq j < N}$, where party S_j has access to an input random variable A_j and generates an output random variable X_j . This allows us to causally order parties: If A_j is correlated to X_k , then S_j is in the causal past of S_k ($S_j \preceq S_k$). To simplify the presentation, we write $\vec{X} = (X_0, \dots, X_{N-1})$ and likewise for \vec{A} , \vec{O} , and \vec{I} .

Definition 2 (Two-party predefined causal order). A *two-party predefined causal order* is a causal ordering of party S with input A , output X , and party T with input B , output Y , such that the distribution $P_{X,Y|A,B}$ can be written as a convex combination of one-way signaling distributions

$$P_{X,Y|A,B} = pP_{X|A,B,Y}P_{Y|B} + (1-p)P_{X|A}P_{Y|A,B,X},$$

for some $0 \leq p \leq 1$.

A definition for multi-party predefined causal order is given in Ref. [42]. Such a definition turns out to be more subtle since a party S_j in the causal past of some other parties $\{S_\ell\}_L$ can in principle influence everything in her causal future; in particular, S_j can influence the causal order of the parties $\{S_\ell\}_L$. We just state a Lemma that follows from such a definition and that is sufficient to prove our claims.

Lemma 1 (Necessary condition for predefined causal order). *A necessary condition for a predefined causal order is that the probability distribution $P_{\vec{X}|\vec{I}}$ can be written as a convex combination*

$$P_{\vec{X}|\vec{I}} = \sum_k p_k P_k,$$

with $\sum_k p_k = 1$ and $\forall k : p_k \geq 0$, such that in every distribution P_k at least one party is not in the causal future of any other party, *i.e.*,

$$\forall k \exists i \forall j \neq i : S_i \not\preceq^{P_k} S_j,$$

where $\not\preceq^{P_k}$ stands for the causal relation that is deduced from the distribution P_k .

In the framework without predefined causal order, each party S_j receives a random variable I_j from the environment E on which S_j can act. After the interaction with I_j , party S_j outputs a random variable O_j to the environment. Both random variables I_j and O_j are *output* random variables. The only *input* random variable a party has is A_j . The operation of S_j is a stochastic process mapping A_j, I_j to X_j, O_j (see Figure 3). A stochastic process is a probability distribution over the range conditioned on the domain; in this case, the stochastic process of party S_j (which in the following will also be called the *local operation* of party S_j) is $P_{X_j, O_j|A_j, I_j}$. All parties are allowed to apply *any* possible operation described by probability theory. Furthermore, they are isolated from each other, which means that they can interact only through the environment. Because we do not make global assumptions (beyond logical consistency), the most general picture is that the random variables that are sent from the environment E to the parties are the result of a map on the random variables fed back by all parties to the same environment E (see Figure 4). Such a composition of parties with the environment combines states and communication channels in one framework.

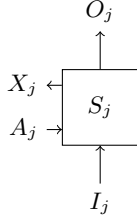


Figure 3. A single party S_j describes a stochastic process $P_{X_j, O_j | A_j, I_j}$. The variables A_j , and X_j model the input and the output. The variable I_j is obtained from the environment E ; the party S_j feeds the variable O_j into the same environment.

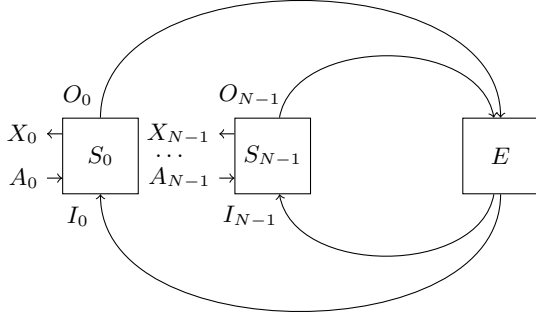


Figure 4. The box E describes the environment. Because no predefined causal order is assumed between the parties, the random variable obtained by the parties is the result of E applied to the outgoing random variable of all parties. This picture combines states and channels, *i.e.*, signaling and non-signaling correlations. For example, assume that S_0 is in the causal past all other parties. In that case, the random variable I_0 is constant, whereas the random variable $I_{j(\neq 0)}$ could depend on A_0 . For three parties or more, this framework gives rise to a new quality: E can describe a map where no I_j is a constant, yet where no contradiction arises. Such correlations are called non-causal. Similarly to the parties, the box E is a stochastic process $P_{I_0, \dots, I_{N-1} | O_0, \dots, O_{N-1}}$.

A party S_j has access to the four random variables X_j , O_j , I_j , and A_j , where A_j is chosen freely. If we consider all parties together, we should get a probability distribution $P_{\vec{X}, \vec{I}, \vec{O} | \vec{A}}$. Furthermore, we ask the environment E to be a multi-linear functional of all local operations. The motivation for this is that linear combinations of local operations should carry through to the probabilities $P_{\vec{X}, \vec{O}, \vec{I} | \vec{A}}$. This brings us to a definition of logical consistency.

Definition 3 (Logical consistency). An environment E is called *logically consistent* if and only if it is a multi-linear positive map on *any* choice of local operations $\{P_{X_j, O_j | A_j, I_j}\}_{0 \leq j < N}$ of all parties such that the composition of E with the local operations results in a probability distribution $P_{\vec{X}, \vec{I}, \vec{O} | \vec{A}}$.

The linearity and positivity conditions from Definition 3 imply Theorem 1, which states that the environment must be a stochastic process (conditional probability distribution).

Theorem 1 (Logical consistent environment as stochastic process). *The environment E is a stochastic process $P_{\vec{I} | \vec{O}}$ that maps \vec{O} to \vec{I} .*

Proof. The environment is a multi-linear positive map \mathcal{E} on the probabilities (we omit the arguments for the sake of presentation)

$$p_j := P_{X_j, O_j | A_j, I_j}(x_j, o_j, a_j, i_j)$$

that party S_j outputs o_j to the environment and generates x_j conditioned on the setting a_j and on $I_j = i_j$. Therefore, we write

$$P_{\vec{X}, \vec{I}, \vec{O} | \vec{A}}(\vec{x}, \vec{i}, \vec{o}, \vec{a}) = \mathcal{E}(p_0, \dots, p_{N-1}).$$

Since \mathcal{E} is a multi-linear positive map and since it depends on \vec{O} and \vec{I} only, the above probability can be written as

$$P_{\vec{X}, \vec{I}, \vec{O} | \vec{A}}(\vec{x}, \vec{i}, \vec{o}, \vec{a}) = E(\vec{o}, \vec{i}) p_0 \cdots p_{N-1}, \quad (1)$$

where $E(\vec{o}, \vec{i})$ is a number. This number must be non-negative, as otherwise the above expression (1) is not a probability. By fixing $\vec{A} = \vec{a}$ and by summing over \vec{x} , we get

$$\begin{aligned} P_{\vec{I}, \vec{O} | \vec{A} = \vec{a}}(\vec{i}, \vec{o}) &= \sum_{\vec{x}} E(\vec{o}, \vec{i}) p_0 \cdots p_{N-1} \\ &= E(\vec{o}, \vec{i}) \sum_{\vec{x}} p_0 \cdots p_{N-1} \\ &= E(\vec{o}, \vec{i}) p'_0 \cdots p'_{N-1}, \end{aligned}$$

where

$$p'_j := P_{O_j | I_j, A_j = a_j}(o_j, i_j).$$

Let us fix the local operations p'_j of all parties to be

$$P_{O_j | I_j, A_j = a_j}(o_j, i_j) = \begin{cases} 1 & o_j = 0, \\ 0 & \text{otherwise.} \end{cases}$$

From the total-probability condition we obtain

$$\sum_{\vec{o}, \vec{i}} E(\vec{o}, \vec{i}) p'_0 \cdots p'_{N-1} = \sum_{\vec{i}} E(\vec{0}, \vec{i}) = 1.$$

By repeating this calculation for different choices of local operations where the parties deterministically output a value, we get

$$\forall \vec{o}: \sum_{\vec{i}} E(\vec{o}, \vec{i}) = 1.$$

Therefore, E is a stochastic process $P_{\vec{I} | \vec{O}}$. □

The following Corollary follows from Theorem 1.

Corollary 1. A logical consistent environment $P_{\bar{I}|\bar{O}}$ fulfills the property that under any choice of the local operations $\{P_{X_j, O_j | A_j, I_j}\}_{0 \leq j < N}$ of all parties, the expression $P_{\bar{I}|\bar{O}} \prod_{j=0}^{N-1} P_{X_j, O_j | A_j, I_j}$ form a conditional probability distribution $P_{\bar{X}, \bar{I}, \bar{O} | \bar{A}}$.

Note that not every conditional distribution $P_{\bar{I}|\bar{O}}$ is logically consistent. Some stochastic processes lead to *grandfather-paradox-type* [50] inconsistencies. Consider the following two extreme examples of such inconsistencies. We describe the examples in the single-party scenario as depicted in Figure 5 and where O , I , X , and A are binary random variables.

Example 1. Let the environment as well as the party S forward the random variable, i.e., the operation of the environment is

$$P_{I|O}(i, o) = \begin{cases} 1 & i = o, \\ 0 & \text{otherwise,} \end{cases}$$

and the operation of the party S is

$$P_{X, O | A, I}(x, o, a, i) = \begin{cases} 1 & o = i = x, \\ 0 & \text{otherwise.} \end{cases}$$

Since the environment E and the party S forward the random variable, we have $\Pr(O = I) = 1$. However, it is unclear what value the probability $P_O(0)$ should take. This is also known as the *causal-loop paradox*.

Example 2. We alter the local operation of party S to *negate* the binary random variable

$$P_{X, O | A, I}(x, o, a, i) = \begin{cases} 1 & o = i \oplus 1 = x, \\ 0 & \text{otherwise.} \end{cases}$$

Now, we are faced with the *grandfather paradox*: if party S receives $i = 1$ from the environment, then she sends the value $o = 0$ to the environment. But in that case, she should receive $i = 0$ and not $i = 1$.

B. Mathematical model of states, operations, evolution, and composition

Let $\{q_0, q_1, \dots\}$ be the sample space of a random variable Q with the probability measure P_Q .

Definition 4 (States, operations, evolution, and composition). We represent a *state* corresponding to a random variable P_Q as the probability vector

$$\vec{P}_Q = (P_Q(q_0), P_Q(q_1), \dots)^T.$$

A stochastic process $P_{R|Q}$ from Q to a random variable R with sample space $\{r_0, r_1, \dots\}$ describes an *operation* and is modeled by the stochastic matrix

$$\hat{P}_{R|Q} = \begin{pmatrix} P_{R|Q}(r_0, q_0) & P_{R|Q}(r_0, q_1) & \dots \\ P_{R|Q}(r_1, q_0) & P_{R|Q}(r_1, q_1) & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

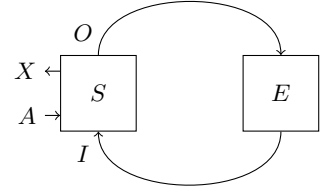


Figure 5. Party S is described by $P_{X, O | A, I}$ and the environment E is $P_{I|O}$.

The result P_R of *evolving* the random variable P_Q through the operation $P_{R|Q}$ is given by the matrix multiplication

$$\vec{P}_R = \hat{P}_{R|Q} \vec{P}_Q.$$

Finally, vectors and matrices are *composed* in parallel using the Kronecker product \otimes .

For example, by this definition, the output of a stochastic process $P_{R|Q_0, Q_1}$ taking two inputs and producing one output is expressed by

$$\hat{P}_{R|Q_0, Q_1} (\vec{P}_{Q_0} \otimes \vec{P}_{Q_1}).$$

C. Set of logically consistent processes without predefined causal order

We derive the conditions on the environment E (stochastic process) such that it is *logically consistent*. For simplicity, we start with the single-party scenario as depicted in Figure 5; the party is denoted by S and the environment by E . We can further simplify our picture by fixing the value of A to a and by summing over X :

$$\sum_x P_{X=x, O | A=a, I} = P_{O|I}.$$

The stochastic process of the environment E is $P_{I|O}$. For now, let us assume that S performs a deterministic operation $D_{O|I}$. This assumption is dropped later. The operation applied by S can be written as a function

$$o = f(i),$$

where i is a deterministic input value. By embedding f into the process of E , we get

$$P_{I|O}(i, f(i)).$$

This can be interpreted as a probability measure of party S receiving the value i from the environment E :

$$Q(i) = P_{I|O}(i, f(i)).$$

For $Q(i)$ to represent a probability measure, the values of Q for every deterministic value i must be non-negative

and have to sum up to 1:

$$\begin{aligned} \forall i : Q(i) &\geq 0, \\ \sum_i Q(i) &= 1. \end{aligned}$$

We express both conditions in the matrix picture. Non-negativity is achieved whenever all entries of the matrix $\hat{P}_{I|O}$ are non-negative. The total-probability conditions are formulated in the following way. The value $f(i)$ that is fed into the environment E is

$$\hat{D}_{O|I=\vec{i}} = \hat{D}_{O|I}\vec{i}.$$

The matrix $\hat{P}_{I|O}$ fixed to providing the state \vec{i} to the party S is

$$\hat{P}_{I=\vec{i}|O} = \vec{i}^T \hat{P}_{I|O}.$$

Therefore, the probability of party S observing i is

$$Q(i) = \vec{i}^T \hat{P}_{I|O} \hat{D}_{O|I}\vec{i},$$

and the law of total probability requires

$$\text{Tr}(\hat{P}_{I|O} \hat{D}_{O|I}) = 1.$$

This condition remains the same if we relax the input to a stochastic input and the operation of S to a stochastic process $P_{O|I}$. The reason for this is that any stochastic input can be written as a convex combination of deterministic inputs, and any stochastic process can be written as a convex combination of deterministic operations. Therefore, the logical-consistency requirement asks the environment E to be restricted to those processes \hat{E} where, under *any choice* of the local operation $P_{O|I}$ of party S , the law of total probability

$$\text{Tr}(\hat{E} \hat{P}_{O|I}) = 1 \quad (2)$$

and the non-negativity condition

$$\forall i, j : \hat{E}_{i,j} \geq 0 \quad (3)$$

hold. Because a stochastic process can be written as a convex mixture of deterministic operations, it is sufficient to ask for

$$\begin{aligned} \forall \hat{D} \in \mathcal{D} : \text{Tr}(\hat{E} \hat{D}) &= 1 \\ \forall i, j : \hat{E}_{i,j} &\geq 0 \end{aligned}$$

for every operation \hat{D} from the set \mathcal{D} of all deterministic operations. Thanks to linearity, we can straightforwardly extend these requirements to multiple parties, and arrive at Theorems 2 and 3.

Theorem 2 (Total probability). *The law that the sum of the probabilities over the exclusive states the parties receive is 1 is satisfied if and only if*

$$\forall \hat{D}_0, \hat{D}_1, \dots \in \mathcal{D} : \text{Tr}(\hat{E}(\hat{D}_0 \otimes \hat{D}_1 \otimes \dots)) = 1,$$

where \hat{D}_j represents a deterministic operation of party S_j .

Theorem 3 (Non-negative probabilities). *The law that the probability of the parties observing a state is non-negative is satisfied if and only if*

$$\forall i, j : \hat{E}_{i,j} \geq 0.$$

D. Equivalence to the quantum correlations framework in the classical limit

The ingredients of the framework by Oreshkov, Costa, and Brukner [31] are process matrices and local operations — described by matrices as well. All the matrices are completely-positive trace-preserving quantum maps in the Choi-Jamiołkowski [51, 52] picture. In the classical limit, the matrices become diagonal in the computational basis [31, 35]. In the single-party scenario, the process matrix W is a map from the Hilbert space \mathcal{H}_O to the Hilbert space \mathcal{H}_I . The party's local operation A then again is a map from the Hilbert space \mathcal{H}_I to the Hilbert space \mathcal{H}_O . The conditions a process matrix W in a single-party scenario has to fulfill [31] are

$$\forall A \in \mathcal{M} : \text{Tr}(WA) = 1, \quad (4)$$

$$W \geq 0, \quad (5)$$

where \mathcal{M} is the set of all completely-positive trace-preserving maps from the space \mathcal{H}_I to the space \mathcal{H}_O . Intuitively, the condition given by Equation (4) “short-circuits” both maps and enforces the probabilities of the outcomes to sum up to 1.

Theorem 4 (Equivalence). *The quantum framework given by Equations (4) and (5) in the classical limit is equivalent to the description of classical correlations without predefined causal order given by Equations (2) and (3).*

Proof. The process matrix W in the quantum framework corresponds to the stochastic process of the environment E in our framework, and the local operations correspond to the stochastic process of the parties. We show a bijection between process matrices and stochastic processes of the environment, and between local operations and stochastic processes of the parties.

A stochastic matrix \hat{E} , representing the environment E in our framework, can be translated into the quantum framework by

$$W_{\hat{E}} = \sum_k |k\rangle\langle k|_{\mathcal{H}_O} \otimes d \left(\hat{E}|k\rangle \sum_{\ell} \langle \ell|_{\mathcal{H}_I} \right),$$

where $|k\rangle$ and $|\ell\rangle$ are computational-basis states of the same dimension as \hat{E} , and where the subscripts denote the respective Hilbert spaces. This completely-positive trace-preserving map (expressed in the Choi-Jamiołkowski picture) acts in the same way as the stochastic matrix \hat{E} : The state $|k\rangle$ is mapped to $\hat{E}|k\rangle$.

The function $d(\rho)$ takes the matrix ρ and cancels all off-diagonal terms, *i.e.*,

$$d(\rho) = \sum_m |m\rangle\langle m|\rho|m\rangle\langle m|.$$

We can rewrite $W_{\hat{E}}$ as

$$W_{\hat{E}} = \sum_k |k\rangle\langle k|_{\mathcal{H}_O} \otimes \sum_m |m\rangle\langle m|\hat{E}|k\rangle\langle m|_{\mathcal{H}_I}.$$

Analogously, the stochastic matrix $\hat{P}_{O|I}$ of the party can be translated into the quantum framework and becomes

$$A_{\hat{P}_{O|I}} = \sum_{k',m'} |m'\rangle\langle m'|\hat{P}_{O|I}|k'\rangle\langle m'|_{\mathcal{H}_O} \otimes |k'\rangle\langle k'|_{\mathcal{H}_I}.$$

The reverse direction of the bijection follows from the description above.

Now, we show that the conditions (4) and (5) in a single-party scenario on a process matrix W coincide with the conditions (2) and (3) in our framework. The non-negativity condition (5) forces the probabilities of the outputs of W to be non-negative; the same holds for the condition (3) in our framework. That the condition (4) coincides with the condition (2) is shown below. Forcing W and A to be diagonal in the computational basis gives

$$\begin{aligned} \text{Tr}(WA) &= \sum_{i,j} \langle i,j|WA|i,j\rangle \\ &= \sum_{i,j} \langle i,j|W|i,j\rangle \langle i,j|A|i,j\rangle. \end{aligned}$$

Substituting W with $W_{\hat{E}}$ and A with $A_{\hat{P}_{O|I}}$ yields

$$\begin{aligned} &\sum_{i,j,m,k,m',k'} \langle i|k\rangle\langle k|i\rangle\langle j|m\rangle\langle m|\hat{E}|k\rangle\langle m|j\rangle \times \\ &\quad \langle i|m'\rangle\langle m'|\hat{P}_{O|I}|k'\rangle\langle m'|i\rangle\langle j|k'\rangle\langle k'|j\rangle \\ &= \sum_{i,j} \langle j|\hat{E}|i\rangle\langle i|\hat{P}_{O|I}|j\rangle \\ &= \text{Tr}(\hat{E}\hat{P}_{O|I}), \end{aligned}$$

which proves the claim. The multi-party case follows through linearity. \square

III. POLYTOPE OF CLASSICAL PROCESSES WITHOUT PREDEFINED CAUSAL ORDER

A. Polytopes

Convex polytopes can be represented in two different ways: The *H-representation* is a list of half-spaces where the intersection is the polytope, and the *V-representation* is a list of the extremal points of the polytope. Algorithms like the double-description method [53, 54] enumerate all extremal points of the polytope given the

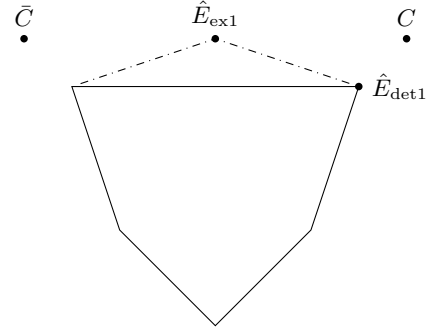


Figure 6. Here, we see a projection of the polytope of classical processes without predefined causal order among three parties and with binary inputs/outputs. The circular identity channel C and the circular bit-flip channel \bar{C} are logically inconsistent; they can be used to reproduce the grandfather's paradox. The solid lines mark the deterministic-extrema polytope and the dashed-dotted lines mark the additional space of logically consistent processes. Point \hat{E}_{ex1} is an extremal point of the polytope and is a uniform mixture of the deterministic processes C and \bar{C} . The behavior of this point is shown in Figure 12. Point \hat{E}_{det1} is an extremal point of the deterministic-extrema polytope, and is described in Figure 11.

H-representation. We used cdd+ [55] for vertex enumeration. The inverse problem is solved by its dual: a convex-hull algorithm.

Here, we derive the polytope of classical processes without predefined causal order. This polytope is represented by the dashed-dotted lines in Figure 2. A projection of the polytope for three parties and binary inputs/outputs onto a plane is given in Figure 6.

B. Single party, binary input, and binary output

We start with the polytope for one party (see Figure 5) with a binary input and a binary output. In this case, a process is described by a square matrix of dimension 2. The most general process of the environment E is

$$\hat{E} = \hat{P}_{I|O} = \begin{pmatrix} w_0 & w_1 \\ w_2 & w_3 \end{pmatrix},$$

consisting of 4 variables. The deterministic operations party S can apply are

$$\begin{aligned} \hat{D}_0 &= \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, & \hat{D}_1 &= \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \\ \hat{D}_2 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & \hat{D}_3 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \end{aligned}$$

where \hat{D}_0, \hat{D}_1 produce a constant 0, 1, respectively, and where the matrix \hat{D}_2 is the identity and \hat{D}_3 the negation.

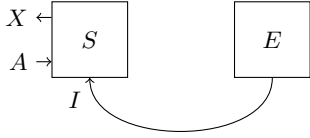


Figure 7. All logically consistent single-party correlations that can be obtained with a feedback channel (see Figure 5) can be simulated without feedback channel.

The equalities

$$\text{Tr}(\hat{E}\hat{D}_0) = 1, \quad (6)$$

$$\text{Tr}(\hat{E}\hat{D}_1) = 1, \quad (7)$$

$$\text{Tr}(\hat{E}\hat{D}_2) = 1, \quad (8)$$

enforce

$$\text{Tr}(\hat{E}\hat{D}_3) = 1.$$

This is shown as follows:

$$\begin{aligned} & \text{Tr}(\hat{E}\hat{D}_0) + \text{Tr}(\hat{E}\hat{D}_1) + \text{Tr}(\hat{E}\hat{D}_2) \\ &= (w_0 + w_2) + (w_1 + w_3) + (w_0 + w_3) \\ &= 2(w_0 + w_3) + w_1 + w_2 \\ &= 2\text{Tr}(\hat{E}\hat{D}_2) + \text{Tr}(\hat{E}\hat{D}_3). \end{aligned}$$

By eliminating three variables using the total-probability conditions (6), (7), and (8) from above, we get

$$\hat{P}_{I|O} = \begin{pmatrix} w_0 & w_0 \\ 1 - w_0 & 1 - w_0 \end{pmatrix}$$

with the non-negativity conditions

$$\begin{aligned} w_0 &\geq 0, \\ 1 - w_0 &\geq 0. \end{aligned}$$

This solution set is a one-dimensional polytope with the extremal points 0 and 1. All solutions describe a state. This implies that all correlations that can be obtained in this framework with a single party and binary input and output, can also be obtained in a framework without feedback, *i.e.*, these correlations can be obtained causally (see Figure 7).

C. Two parties, binary inputs, and binary outputs

In the two-party case with a binary input and a binary output for each party, the process $\hat{E} = \hat{P}_{I_0, I_1 | O_0, O_1}$ of the environment is described by a square matrix of dimension 2^2 . The conditions are

$$\begin{aligned} \forall i, j \in \{0, 1, 2\} : \text{Tr}(\hat{E}(\hat{D}_i \otimes \hat{D}_j)) &= 1, \\ \forall i, j : \hat{E}_{i,j} &\geq 0. \end{aligned} \quad (9)$$

With a similar argument as above, one can show that the operation \hat{D}_3 does not need to be considered for either party. The matrix \hat{E} consists of 4^2 unknowns, out of which 3^2 are eliminated by the total-probability conditions given by Equation (9). Thus, we are left with 7 unknowns, forming a 7-dimensional polytope with 16 inequalities.

The resulting V -representation of the polytope consists of 12 extremal points, all of which represent deterministic processes:

$$\begin{aligned} \hat{E}_0 &= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & \hat{E}_1 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ \hat{E}_2 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & \hat{E}_3 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \\ \hat{E}_4 &= \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & \hat{E}_5 &= \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ \hat{E}_6 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}, & \hat{E}_7 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}, \\ \hat{E}_8 &= \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & \hat{E}_9 &= \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ \hat{E}_{10} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}, & \hat{E}_{11} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}. \end{aligned}$$

In the following, we use $A = S_0$ and $B = S_1$. The first four processes $\hat{E}_0, \hat{E}_1, \hat{E}_2, \hat{E}_3$ represent the four constants (0,0), (0,1), (1,0), (1,1) as inputs to the parties A and B . The next four processes represent a constant input to party A (processes \hat{E}_4 and \hat{E}_5 produce the constant 0, and the other two processes produce the constant 1) and a channel from party A to party B ; the processes \hat{E}_4 and \hat{E}_6 describe the identity channel, and \hat{E}_5 and \hat{E}_7 describe the bit-flip channel. The last four processes are analogous, with a channel from B to A and where party B receives a constant. All these 12 processes act deterministically on bits for two parties where at least one party receives a constant (see Figure 8). Therefore, every such channel can be simulated in a causal fashion. This result generalized to higher dimensions was already shown by taking the classical limit of the framework for quantum correlations without predefined causal order [31].

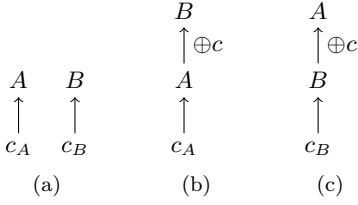


Figure 8. (a) Both parties A and B receive a constant each. (b) Party A receives a constant and sends a bit through the identity ($c = 0$) or the bit-flip ($c = 1$) channel to B . (c) Same as (b), where the parties are interchanged.

D. Three parties, binary inputs, and binary outputs

The process of the environment E in a three party setup with binary inputs and outputs is described by a square matrix $\hat{E} = \hat{P}_{I_0 I_1 I_2 | O_0 O_1 O_2}$ of dimension 2^3 . The matrix \hat{E} consists of 4^3 variables, out of which 3^3 can be eliminated with the total-probability conditions

$$\forall i, j, k \in \{0, 1, 2\} : \text{Tr} \left(\hat{E} \left(\hat{D}_i \otimes \hat{D}_j \otimes \hat{D}_k \right) \right) = 1, \quad (10)$$

resulting in a 37-dimensional polytope with 4^3 linear constraints (non-negative probabilities):

$$\forall i, j : \hat{E}_{i,j} \geq 0. \quad (11)$$

Solving this polytope yields 710'760 extremal points. Only 744 extremal points out of these 710'760 are deterministic, *i.e.*, consist of 0-1 values; the remaining extremal points are so-called *proper mixtures* of logically inconsistent processes. Such proper mixtures are *not* convex combinations of deterministic extremal points *inside* the polytope, but are convex combinations of deterministic points where some lie *outside* of the polytope — any process from outside of the polytope leads to logical inconsistencies. Interestingly, this smaller polytope (hence, also the polytope described by the Equations (10) and (11)) consists of processes that cannot be simulated using a predefined causal order, *i.e.*, processes where *no party receives a constant*, implying that *every party causally succeeds some other party*. The 744 deterministic extremal points are discussed in Section IV along with the general polytope restricted to the deterministic extremal points.

E. General case

We describe the polytope for logically consistent classical processes without predefined causal order in the general case. Let n be the number of parties and let d be the dimension of the states entering and leaving every laboratory. This leaves us with a $d^n \times d^n$ stochastic matrix \hat{E} describing the environment. Every party can perform an

operation that is a convex mixture of all d^d deterministic operations. The set of all deterministic operations is denoted by \mathcal{D} . For every party, under any choice of deterministic operation $D \in \mathcal{D}$, the trace of the environment \hat{E} multiplied with the local operations is constrained to give 1 (see Theorem 2). However — as in the binary-input/output case above —, some of these constraints are redundant.

Theorem 5 (Sufficient set for total-probability conditions). *The total-probability conditions to this family of operations*

$$\hat{D}_{i,j} = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 0 & 1 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix},$$

where j is output for input i and 0 otherwise, imply the total-probability conditions for all remaining deterministic operations of the same dimension, *i.e.*,

$$\forall i_0, j_0, i_1, j_1, \dots, i_{n-1}, j_{n-1} \geq 0 :$$

$$\text{Tr} \left(\hat{E} \left(\hat{D}_{i_0, j_0} \otimes \dots \otimes \hat{D}_{i_{n-1}, j_{n-1}} \right) \right) = 1$$

\implies

$$\forall \hat{D}_0, \dots, \hat{D}_{n-1} \in \mathcal{D} :$$

$$\text{Tr} \left(\hat{E} \left(\hat{D}_0 \otimes \dots \otimes \hat{D}_{n-1} \right) \right) = 1.$$

Proof. We restrict ourselves to the single-party scenario — the multi-party case follows through linearity. Let \vec{v}_i be the d -dimensional vector with a 1-entry at position i and 0's everywhere else. We can write a d -dimensional matrix $\hat{D}_{i,j}$ as

$$\hat{D}_{i,j} = \left(\sum_{m \neq i} \vec{v}_0 \vec{v}_m^T \right) + \vec{v}_j \vec{v}_i^T.$$

A general deterministic matrix $D \in \mathcal{D}$ of the same dimension, where k is mapped to a_k , is expressed as

$$D = \sum_k \vec{v}_{a_k} \vec{v}_k^T.$$

On the one hand, using the antecedent above, we have

$$\text{Tr} \left(\hat{E} \sum_{k: a_k \neq 0} \hat{D}_{k, a_k} \right) = \sum_{k: a_k \neq 0} \text{Tr} \left(\hat{E} \hat{D}_{k, a_k} \right) = \ell \quad (12)$$

with $\ell = |\{k | a_k \neq 0\}|$. On the other hand, we can

rewrite $\sum_{k:a_k \neq 0} \hat{D}_{k,a_k}$ as

$$\begin{aligned}
\sum_{k:a_k \neq 0} \hat{D}_{k,a_k} &= \sum_{k:a_k \neq 0} \left(\sum_{m \neq k} \vec{v}_0 \vec{v}_m^T \right) + \vec{v}_{a_k} \vec{v}_k^T \\
&= \sum_{k:a_k \neq 0} \vec{v}_{a_k} \vec{v}_k^T + \sum_{k:a_k=0} \vec{v}_0 \vec{v}_k^T \\
&\quad - \sum_{k:a_k=0} \vec{v}_0 \vec{v}_k^T + \sum_{k:a_k \neq 0} \sum_{m \neq k} \vec{v}_0 \vec{v}_m^T \\
&= D + \sum_{k:a_k \neq 0} \sum_{m \neq k} \vec{v}_0 \vec{v}_m^T - \sum_{k:a_k=0} \vec{v}_0 \vec{v}_k^T \\
&= D + \ell \sum_{k:a_k=0} \vec{v}_0 \vec{v}_k^T + (\ell - 1) \sum_{k:a_k \neq 0} \vec{v}_0 \vec{v}_k^T \\
&\quad - \sum_{k:a_k=0} \vec{v}_0 \vec{v}_k^T \\
&= D + (\ell - 1) \sum_k \vec{v}_0 \vec{v}_k^T \\
&= D + (\ell - 1) \hat{D}_{0,0}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\text{Tr} \left(\hat{E} \sum_{k:a_k \neq 0} \hat{D}_{k,a_k} \right) &= \text{Tr}(\hat{E}D) + (\ell - 1) \text{Tr}(\hat{E}\hat{D}_{0,0}) \\
&= \text{Tr}(\hat{E}D) + \ell - 1,
\end{aligned}$$

which, with the identity (12), implies

$$\text{Tr}(\hat{E}D) = 1.$$

□

The family $\{\hat{D}_{i,j} | i, j \in I\}$ of deterministic operations with the set $I = \{0, \dots, n-1\}$ has size $d(d-1) + 1$.

Theorem 6 (Polytope). *The H -representation of the polytope of logically consistent classical processes without predefined causal order is*

$$\begin{aligned}
\forall \hat{D}_0, \hat{D}_1, \dots, \hat{D}_{n-1} \in \{\hat{D}_{i,j}\}_{I \times I} : \\
\text{Tr} \left(\hat{E} \left(\hat{D}_0 \otimes \hat{D}_1 \otimes \dots \otimes \hat{D}_{n-1} \right) \right) &= 1, \\
\forall i, j : \hat{E}_{i,j} &\geq 0.
\end{aligned}$$

The polytope has d^{2n} facets and dimension

$$d^{2n} - (d(d-1) + 1)^n,$$

which is exponential in the number of parties.

IV. THE DETERMINISTIC-EXTREMA POLYTOPE

Definition 5 (Deterministic-extrema polytope). The *deterministic-extrema polytope* is defined as the polytope of logically consistent processes without predefined causal order where all extremal points are deterministic processes (see polytope with the solid lines in Figure 2).

The deterministic-extrema polytope excludes *proper mixtures* of logically inconsistent processes. Such mixtures (consistent mixture of inconsistent points) are convex combinations of deterministic points where at least one deterministic point is outside of the polytope. To find this polytope, one can first solve the extremal points of the general polytope and thereafter select the boolean solutions. These boolean solutions form the V -representation of the polytope in discussion.

A. Three parties, binary inputs, and binary outputs

We discuss the deterministic-extrema polytope in the setting of three parties and binary inputs and outputs. To simplify the presentation, we use $A = S_0$, $B = S_1$, $C = S_2$, $O_A = O_0$, $O_B = O_1$, $O_C = O_2$, $I_A = I_0$, $I_B = I_1$, and $I_C = I_2$. As described in Section III D, this polytope has 744 extremal points. They can be characterized as follows.

Assume that \hat{E} , when the parties locally apply the identity operation, maps $(0, 0, 0)$ to $(0, 0, 0)$, *i.e.*, $(0, 0, 0)$ is a fixed-point. Then, any other extremal point \hat{E}' is obtained by the local operations identity \hat{D}_3 and bit-flip \hat{D}_4 , where we embed these local operations into the environment. Let $\hat{L}_{i,j,k}$ be the local operation of the three parties

$$\hat{L}_{i,j,k} = \hat{D}_4^i \otimes \hat{D}_4^j \otimes \hat{D}_4^k,$$

i.e., party A performs the identity if $i = 0$ and the bit-flip operation if $i = 1$ — the other parties' local operations are defined in the same way. The extremal point \hat{E}' can be described as

$$\hat{E}' = \hat{L}_{i,j,k} \hat{E} \hat{L}_{i,j,k},$$

where, as described above, the operations are embedded into the environment (see Figure 9). Logical consistency of the environment \hat{E}' follows because we started with a logically consistent \hat{E} and the operations act on single parties. The process \hat{E}' maps (i, j, k) to (i, j, k) . Thus, starting with \hat{E} , for any choice of i, j, k , we obtain a different extremal point. There are $2^3 - 1$ alternative extremal points that can be constructed in this fashion. From this we conclude that $744/8 = 93$ extremal points are such, that $(0, 0, 0)$ is a fixed-point under locally applying the identity. We restrict our analysis to these 93 extremal points; all others can be obtained by the above construction. The following analysis is structured depending on the number of parties that receive a constant from the environment.

There exists only one extremal point where $(0, 0, 0)$ is mapped to $(0, 0, 0)$ under applying identity locally, and where *every party* receives a constant: the constant $(0, 0, 0)$.

Assume exactly *two parties* receive a constant, which leaves us with three possibilities of choosing them. Fix these parties to be A and B . The third party C receives

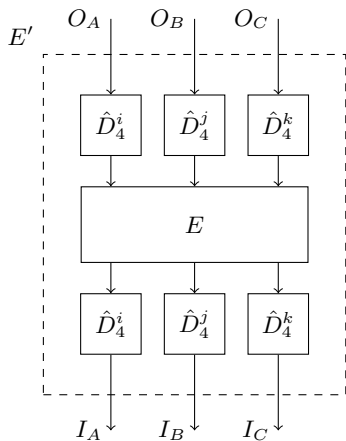


Figure 9. By starting from a logically consistent environment E and for any choice of $i, j, k \in \{0, 1\}$, one can construct another logically consistent environment E' .

a value that depends on the operation of A or of B or of both. Thus, we are in the case $A \preceq C$ or $B \preceq C$. Inevitably, the constant must be $(0, 0)$; otherwise the fixed-point $(0, 0, 0)$ is not recovered. Party C receives a value that depends on the value fed back by A and B ; there exist $2^3 - 1 = 7$ such functions where we have excluded the constant and all operations where C receives a value different from 0 on inputs $(0, 0)$ to the environment from A and B . Therefore, under all permutations of the parties, 21 extremal points give a constant to two parties and have the fixed-point $(0, 0, 0)$ when the identity is applied locally.

In a next step, assume that exactly *one party* receives a constant. This assumption, again, allows for three different setups, as we can choose which party receives a constant. Without loss of generality, let A be this party, *i.e.*, $A \preceq B$ and $A \preceq C$; the constant must be 0 again in order to comply with the requirement of the fixed-point. Now, we are left with several possibilities on how B and C depend on A and on each other. As a first case, we assume that B and C do not depend on each other, but depend on A only ($A \preceq B$ and $A \preceq C$). This dependency cannot be different from the identity channels from A to B and from A to C ; the alternative would be the bit-flip channel that would not reproduce the desired fixed-point $(0, 0, 0)$. This gives us 3 different extremal points under all permutations of the parties. Another possibility on the dependencies is that B depends on A , and C depends on B , *i.e.*, $A \preceq B \preceq C$, and the interchange of parties B and C . The channels — by following the same reasoning above — again must be the identity channels: This gives us 6 extremal points. Now, we look at the case where B depends on A and where C depends on both, A and B , *i.e.*, $A \preceq B$, $A \preceq C$, and $B \preceq C$, and any permutation of the parties. There are 6 permutations. The constant that A receives must be 0, party B must depend trivially on A (the identity channel) and party C can depend in 5 different ways on A and B : These are

all 2-to-1-bit functions where $(0, 0)$ is mapped to 0 (2^3 minus the constant and minus the dependencies on A only and on B only. In total, there are $6 \cdot 5 = 30$ such extremal points. We are left with the last scenario: B depends on A and on C , and C depends on A and on B , *i.e.*, $A \preceq B$, $C \preceq B$, $A \preceq C$, and $B \preceq C$. The constant, as above, is 0. Given the random variable O_A fed to the environment by A , the environment can *either* describe a channel from B to C ($B \preceq C$) or describe a channel from C to B ($C \preceq B$); any other channel would lead to a causal loop. The direction of the channel must differ under different values o_A fed-back by A , as otherwise B and C would not mutually depend on each other. We have two possibilities on the direction given the value fed-back by A is $O_A = 0$. Assume the direction to be $B \preceq C$. For the case $O_A = 0$, the channels from A to B and from B to C are the identity channels in order to comply with the fixed point $(0, 0, 0)$. In the other case $O_A = 1$, the direction of the channel between B and C is in the reverse direction compared to $O_A = 0$, *i.e.*, $C \preceq B$. Then, because of $O_A \neq 0$, the random variables I_B and I_C are not forced to be $(0, 0)$; there exist two channels from A to C and another two channels from B to C . Therefore, we are left with $3 \cdot 2 \cdot 4 = 24$ possibilities. An overview over these setups is given in Figure 10.

The last setup (see Figure 10g) where *no party* receives a constant builds a family of 8 extremal points. All extremal points are equivalent up to relabelling of the inputs to and outputs from the environment. One such extremal point is

$$\hat{E}_{\text{det1}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The behavior of this extremal point is

$$I_A = \bar{O}_B O_C, \quad I_B = O_A \bar{O}_C, \quad I_C = \bar{O}_A O_B, \quad (13)$$

where \bar{x} is the negation of x . This solution is depicted in Figure 11. Its location in the polytope is shown in Figure 6.

V. CAUSAL GAMES

Given a deterministic extremal point where every party is in the causal past of all other parties, a causal game can be constructed that can be won perfectly in the framework presented here, but that is lost if one assumes a global time.

Definition 6 (Causal game). Let a deterministic process map $O = (O_0, O_1, \dots)$ to $I^O = (I_0^O, I_1^O, \dots)$, where

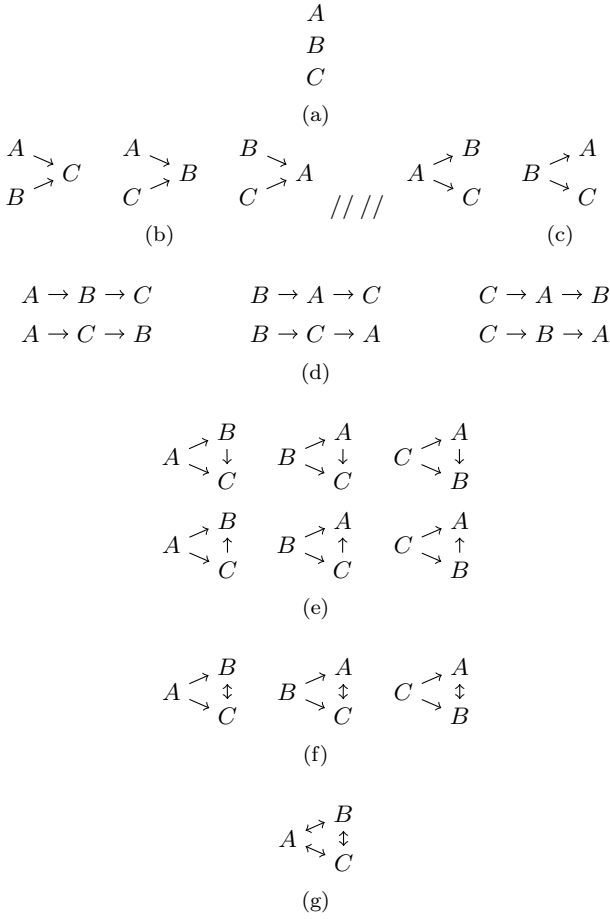


Figure 10. (a) Every party receives a constant. For an environment with fixed point $(0, 0, 0)$ when the parties locally apply the identity map, this constant must be $(0, 0, 0)$. (b) Two parties receive a constant $(0, 0)$, the third party receives a value depending on the other parties' state fed to the environment. For each of the three cases, seven different functions exist. (c) For each of the three cases, the identity function only is consistent with the setup. (d) Here as well, only the identity channel is consistent with the fixed-point. (e) Five different functions are possible per setup. (f) Here, eight functions per setup are consistent with the fixed-point. (g) No party receives a constant, yet no contradiction arises under any choice of local operations. For a fixed-point $(0, 0, 0)$ where the parties locally apply the identity map, eight different functions that fulfill these requirements exist. In total, the number of deterministic extremal points where $(0, 0, 0)$ is mapped to $(0, 0, 0)$ when the parties apply the identity operation is 1 (a) $+ 3 \cdot 7$ (b) $+ 3$ (c) $+ 6$ (d) $+ 6 \cdot 5$ (e) $+ 3 \cdot 8$ (f) $+ 8$ (g) $= 93$.

the i -th entry belongs to party S_i , and where for every i , I_i^O depends on all other parties' inputs to the environment. We define a *causal game* where party S_i gets a random A_i and has to produce $X_i = I_i^A$. The parties are allowed to communicate in a predefined causal order [32, 35, 42]. Let the guesses of all parties be $X = (X_0, X_1, \dots)$, and let the random inputs to all parties be $A = (A_0, A_1, \dots)$. In a setup with n parties and where every party obtains and sends a d -dimensional

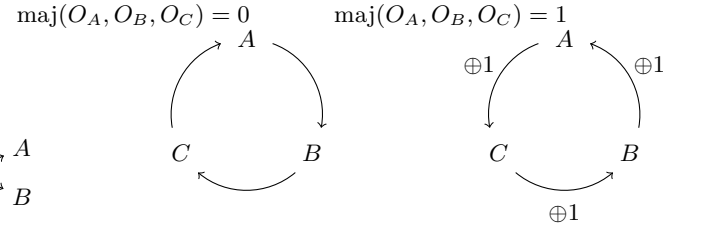


Figure 11. The left channel is chosen if the majority of the values fed into the environment is 0, otherwise, the right channel is chosen (see Equation 13).

state, the game's winning probability is

$$p_{\text{succ}} = \frac{1}{d^n} \sum_I \Pr(X = I^A | A). \quad (14)$$

Let p_{succ}^C , $p_{\text{succ}}^{\text{NC}}$ be the success probability of the game (14) with, without the assumption of a predefined causal order, respectively.

Theorem 7 (No winning strategy with predefined order). *Using a predefined causal order, the success probability (14) is strictly less than 1, i.e., $p_{\text{succ}}^C < 1$.*

Proof. For every party S_i , the random variable I_i^A party S_i has to guess *depends* on the other parties' inputs $A_{j(\neq i)}$. In a predefined causal order, however, at least one party is *not in the causal future of any other party*. Let S_i be that party, i.e., $\forall j \neq i : S_i \not\prec S_j$ (see Lemma 1). Then, at least for one input $A_i = a'$ to S_i , the party S_i cannot guess perfectly. The success probability p_{succ}^C is upper bounded by

$$\begin{aligned} p_{\text{succ}}^C &= \frac{1}{d^n} \left(\sum_{A, A_i \neq a'} \Pr(X = I^A | A) \right. \\ &\quad \left. + \sum_{A, A_i = a'} \Pr(X = I^A | A) \right) \\ &\leq \frac{1}{d^n} \left((d-1)d^{n-1} + \sum_{A, A_i = a'} \Pr(X = I^A | A) \right). \end{aligned}$$

The guessing probability for the non-perfect guess is upper bounded by

$$\Pr(X = I^A | A, A_i = a') \leq \frac{d^{n-1} - 1}{d^{n-1}}$$

because for at least one input A with $A_i = a'$, party S_i guesses wrongly. Therefore, we obtain

$$\begin{aligned} p_{\text{succ}}^C &\leq \frac{1}{d^n} \left((d-1)d^{n-1} + d^{n-1} \frac{d^{n-1} - 1}{d^{n-1}} \right) \\ &= \frac{1}{d^n} (d^n - d^{n-1} + d^{n-1} - 1) \\ &= 1 - \frac{1}{d^n}. \end{aligned}$$

□

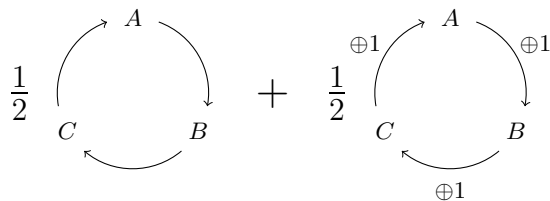


Figure 12. Channel where the circular identity channel is uniformly mixed with the circular bit-flip channel.

Theorem 8 (Winning strategy without predefined causal order). *If we drop the assumption of a predefined causal order, then the causal game (14) can be won perfectly, i.e., $p_{\text{succ}}^{NC} = 1$.*

Proof. To win the game perfectly, the parties use the process that maps the random variable $O = (O_0, O_1, \dots)$ to the random variable $I^O = (I_0^O, I_1^O, \dots)$ deterministically, forward their random input to the environment ($A_i = O_i$), and use the value obtained from the environment as their guess ($X_i = I_i^O$). \square

For other games, a larger gap between the success probability with a predefined causal order and the success probability without a predefined causal order can be achieved — as is shown in the examples below.

VI. EXAMPLES

We briefly discuss two examples in the three-party scenario. Let A, B, C be random input bits to the three parties A, B, C , respectively, and let X, Y, Z be the corresponding output bits.

Example 3. An extremal point of the first class of polytopes for three parties and binary inputs/outputs is

$$\hat{E}_{\text{ex1}} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix};$$

its behavior is shown in Figure 12. This extremal point is a *proper mixture* of logically inconsistent processes, as it cannot be written as a convex combination of deterministic points from within the polytope; the left and the right channels from Figure 12 individually describe a causal loop and, hence, are logically inconsistent. Initially, this process was used to show that in the classical scenario with three parties or more, correlations without predefined causal order can arise [35]. A causal game

that can be formulated for this extremal point is

$$p_{\text{succ}}^{\text{ex1}} = \frac{1}{3} (\Pr(X = B \oplus C \mid m = 1) + \Pr(Y = A \oplus C \mid m = 2) + \Pr(Z = A \oplus B \mid m = 3)),$$

where, depending on the shared random trit m , the party selected by m has to guess the parity of the other two parties' inputs. If we assume a predefined causal order, then this causal game can be won with probability at most $5/6$ [32]. The reason for this is that at least one party is not in the causal future of the others. This party, hence, can guess the parity with a probability of $1/2$ only. However, by using the environment from Figure 12, one can win the game perfectly. To achieve this, if $m = 1$, then party B forwards the random input ($O_B = B$), party C forwards the parity of the random input and the random variable obtained from the environment ($O_C = C \oplus I_C$), and party A uses the random variable obtained from the environment as its guess ($X = I_A$). For the cases $m = 2$ and $m = 3$, the same strategy is used, but where the parties are permuted.

Example 4. Another example [56] is depicted in Figure 11 and is a deterministic extremal point of the polytope with three parties and binary inputs/outputs (see also Equation (13)). Consider the causal game

$$p_{\text{succ}}^{\text{ex2}} = \frac{1}{2} (\Pr(X = C, Y = A, Z = B \mid \text{maj}(A, B, C) = 0) + \Pr(X = \bar{B}, Y = \bar{C}, Z = \bar{A} \mid \text{maj}(A, B, C) = 1)),$$

where $\text{maj}(A, B, C)$ is the majority of the three bits A, B , and C . Whenever the majority of the inputs is 0, i.e., $\text{maj}(A, B, C) = 0$, then the parties play the “guess-your-neighbours-input” game [57, 58]: Party A guesses the input of party B , party B guesses the input of party C , and finally party C guesses the input of party A ; the game is won if all guesses are correct simultaneously. If the majority of the inputs is 1, then they play the same game in reverse direction and flip the output bits. The success probability of winning this game in a world with a predefined causal order is upper bounded by $3/4$. This can be seen by the following reasoning. In a predefined causal order, at least a single party has to make a guess without learning anything from the other parties. For example, if party A causally precedes B and C , i.e., $(A \preceq B \text{ and } A \preceq C)$, then party A at best always outputs 0 (see Table I). By making such a guess, however, in 2 out of 8 cases, the parties will lose the game, yielding an upper bound of $3/4$ to the success probability. The same upper bound is achieved by choosing party B or party C as causally preceding the others.

By using the environment shown in Figure 11, the game can be won perfectly. The parties simply forward their inputs to the environment and use the bits obtained from the environment as the guesses.

A	B	C	$\text{maj}(A, B, C)$	X	Y	Z
0	0	0	0	0	0	0
0	0	1	0	1	0	0
0	1	0	0	0	0	1
0	1	1	1	0	0	1
1	0	0	0	0	1	0
1	0	1	1	1	0	0
1	1	0	1	0	1	0
1	1	1	1	0	0	0

Table I. Conditions for winning the game of Example 4.

VII. CONCLUSION AND OPEN QUESTIONS

We describe the polytope for classical multi-party processes without predefined causal order but where the arising correlations are logically consistent. We also describe the polytope formed by *deterministic* extremal points; it excludes processes that are *proper mixtures* of logically inconsistent processes, *i.e.*, processes that cannot be written as a convex combination of deterministic processes from *within* the polytope. For three parties or more, these polytopes contain processes that cannot be simulated by using a predefined causal order among the parties — this is shown by violations of so-called causal

inequalities.

A representation with polytopes helps for finding new causal games as well as for optimizing the processes for winning causal games; the optimization problem can be stated as a linear program.

In comparison, it has been shown that the set of *causal* correlations, *i.e.*, correlations *with* predefined causal order, also forms a polytope [42, 43]. A complete characterization for the two-party case is given [43], however, in the multi-party case, a characterization is missing. Such a characterization is interesting as then one could subtract it from the polytope studied in this work; this yields an exact characterization of the *non-causal* processes. Another open question is to decide for which causal games the quantum correlations without causal order outperform their classical counterpart.

ACKNOWLEDGMENTS

We thank Mateus Araújo, Veronika Baumann, Cyril Branciard, Časlav Brukner, Fabio Costa, Adrien Feix, Arne Hansen, Alberto Montina, and Benno Salwey for helpful discussions. Furthermore we thank the anonymous referees for the detailed comments. The present work was supported by the Swiss National Science Foundation (SNF) and the National Centre of Competence in Research “Quantum Science and Technology” (QSIT).

-
- [1] L. Hardy, Journal of Physics A: Mathematical and Theoretical **40**, 3081 (2007).
 - [2] A. Einstein, B. Podolsky, and N. Rosen, Physical Review **47**, 777 (1935).
 - [3] J. S. Bell, Physics **1**, 195 (1964).
 - [4] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Reviews of Modern Physics **86**, 419 (2014).
 - [5] H. Reichenbach, in *The Direction of Time*, edited by M. Reichenbach (University of California Press, Berkeley, 1956) Chap. 19, pp. 157–167.
 - [6] A. Suarez and V. Scarani, Physics Letters A **232**, 9 (1997).
 - [7] A. Stefanov, H. Zbinden, N. Gisin, and A. Suarez, Physical Review Letters **88**, 120404 (2002).
 - [8] V. Scarani and N. Gisin, Physics Letters A **295**, 167 (2002).
 - [9] S. Coretti, E. Hänggi, and S. Wolf, Physical Review Letters **107**, 100402 (2011).
 - [10] J.-D. Bancal, S. Pironio, A. Acín, Y.-C. Liang, V. Scarani, and N. Gisin, Nature Physics **8**, 867 (2012).
 - [11] T. J. Barnea, J.-D. Bancal, Y.-C. Liang, and N. Gisin, Physical Review A **88**, 022123 (2013).
 - [12] V. Scarani, J.-D. Bancal, A. Suarez, and N. Gisin, Foundations of Physics **44**, 523 (2014).
 - [13] C. J. Wood and R. W. Spekkens, New Journal of Physics **17**, 033002 (2015).
 - [14] G. W. Leibniz and S. Clarke, in *Correspondence*, edited by R. Ariew (Hackett Publishing Company, Inc., Indianapolis, 2000) pp. 14–18.
 - [15] C. F. von Weizsäcker, “Aufbau der Physik,” (Carl Hanser Verlag, München, 1985) p. 522.
 - [16] D. N. Page and W. K. Wootters, Physical Review D **27**, 2885 (1983).
 - [17] W. K. Wootters, International Journal of Theoretical Physics **23**, 701 (1984).
 - [18] L. Bombelli, J. Lee, D. Meyer, and R. D. Sorkin, Physical Review Letters **59**, 521 (1987).
 - [19] G. M. D’Ariano, AIP Conference Proceedings **1363**, 63 (2011).
 - [20] P. Erker, *The Quantum Hourglass*, Master’s thesis, ETH Zürich (2014).
 - [21] V. Vedral, preprint arXiv:1408.6965 [quant-ph] (2014).
 - [22] V. Giovannetti, S. Lloyd, and L. Maccone, Physical Review D **92**, 045033 (2015).
 - [23] S. Ranković, Y.-C. Liang, and R. Renner, preprint arXiv:1506.01373 [quant-ph] (2015).
 - [24] R. Colbeck and R. Renner, Nature Communications **2**, 411 (2011).
 - [25] G. Ghirardi and R. Romano, Foundations of Physics **43**, 881 (2013).
 - [26] L. Hardy, preprint arXiv:0509120 [gr-qc] (2005).
 - [27] L. Hardy, in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle*, The Western Ontario Series in Philosophy of Science, Vol. 73 (Springer Netherlands, Dordrecht, 2009) pp. 379–401.
 - [28] G. Chiribella, G. M. D’Ariano, and P. Perinotti, Physical Review A **80**, 022339 (2009).
 - [29] G. Chiribella, Physical Review A **86**, 040301 (2012).

- [30] T. Colnaghi, G. M. D’Ariano, S. Facchini, and P. Perinotti, *Physics Letters A* **376**, 2940 (2012).
- [31] O. Oreshkov, F. Costa, and Č. Brukner, *Nature Communications* **3**, 1092 (2012).
- [32] Ä. Baumeler and S. Wolf, in *Proceedings of International Symposium on Information Theory* (IEEE, Piscataway, 2014) pp. 526–530.
- [33] G. Chiribella, G. M. D’Ariano, P. Perinotti, and B. Valiron, *Physical Review A* **88**, 022318 (2013).
- [34] F. Costa, *Local and causal structures in quantum theory*, Ph.D. thesis, Universität Wien, Wien (2013).
- [35] Ä. Baumeler, A. Feix, and S. Wolf, *Physical Review A* **90**, 042106 (2014).
- [36] Č. Brukner, *New Journal of Physics* **17**, 083034 (2015).
- [37] I. Ibnouhsein and A. Grinbaum, *Physical Review A* **92**, 042124 (2015).
- [38] I. Ibnouhsein, *Quantum Correlations and causal structures*, Ph.D. thesis, Université Paris-Sud (2014).
- [39] T. Morimae, *Physical Review A* **90**, 010101 (2014).
- [40] O. Oreshkov and N. J. Cerf, preprint arXiv:1406.3829 [quant-ph] (2014).
- [41] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, *New Journal of Physics* **17**, 102001 (2015).
- [42] O. Oreshkov and C. Giarmatzi, preprint arXiv:1506.05449 [quant-ph] (2015).
- [43] C. Branciard, M. Araújo, A. Feix, F. Costa, and Č. Brukner, *New Journal of Physics* **18**, 013008 (2016).
- [44] A. Feix, M. Araújo, and Č. Brukner, *Physical Review A* **92**, 052326 (2015).
- [45] O. Oreshkov and N. J. Cerf, *Nature Physics* **11**, 853 (2015).
- [46] Č. Brukner, *Nature Physics* **10**, 259 (2014).
- [47] Č. Brukner, private communication (2014).
- [48] H. Price, *The British Journal for the Philosophy of Science* **42**, 157 (1991).
- [49] J. Woodward, *Making things happen: A theory of causal explanation* (Oxford University Press, Oxford, 2003).
- [50] R. Barjavel, *Le voyageur imprudent* (Denoël, Paris, 1944).
- [51] A. Jamiołkowski, *Reports on Mathematical Physics* **3**, 275 (1972).
- [52] M.-D. Choi, *Linear Algebra and its Applications* **10**, 285 (1975).
- [53] T. S. Motzkin, H. Raiffa, G. L. Thompson, and R. M. Thrall, in *Contributions to Theory of Games*, Annals of Mathematics Studies, Vol. 2, edited by H. W. Kuhn and A. W. Tucker (Princeton University Press, Princeton, 1953).
- [54] K. Fukuda and A. Prodon, in *Combinatorics and Computer Science*, edited by M. M. Deza, R. Euler, and I. Manoussakis (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005) pp. 91–111.
- [55] K. Fukuda, “cdd+-077a,” http://www.inf.ethz.ch/personal/fukudak/cdd_home/cdd.htm (2003).
- [56] M. Araújo and A. Feix, private communication (2014).
- [57] M. L. Almeida, J.-D. Bancal, N. Brunner, A. Acín, N. Gisin, and S. Pironio, *Physical Review Letters* **104**, 230404 (2010).
- [58] A. Acín, M. L. Almeida, R. Augusiak, and N. Brunner, in *Quantum Theory: Informational Foundations and Foils*, edited by G. Chiribella and R. W. Spekkens (Springer Netherlands, Dordrecht, 2015) pp. 465–496.